SI Materials and Methods

GENIST algorithm
The MATLAB source code for GENIST is publically available at https://github.com/madeluis/GENIST.

Step 1. Clustering:
Input: Spatial gene expression data for all genes
Output: Input genes divided in *c* clusters

Clustering is implemented using linkage clustering. The specification of the number of clusters *c*, automatically implemented, is given by the Silhouette index (1). The Silhouette index choses the best number of clusters among a range, $\lfloor n_g/10 \pm 5 \rfloor$, heuristically determined to ensure that a small number of nodes is provided to the inference step (≤~10 genes).

Step 2. GRN inference:
Step 2.1. Inferring intra-cluster connections
Input 1: Time-series gene expression data for all genes in cluster C_n
Input 2: User specification to consider whether genes can be activated by regulators (factors that were activated in the time point prior activation of the gene), by co-regulators (factors that are activated in the same time point of the activation of the gene), or by both
Output: Sub-network for cluster C_n

For each cluster C_n, for $n \in [1,c]$, do:

Step 2.1.1 Selection of potential regulators and co- regulators
We limited the potential regulators of each gene (2):
A gene g_r is a potential regulator of a target gene g_s (denoted $g_r \rightharpoonup g_s$) if and only if it exhibits a ±0.1 × g_r change of expression immediately prior a change of expression of g_s of ±0.1 × g_s:

$$g_r \rightharpoonup g_s \leftrightarrow (g_r(t) > 1.1{\times}g_r(t-1) | g_r(t) < 0.9{\times}g_r(t-1)) \,\&\, \left(g_s(t+1) > 1.1 \times g_s(t) \middle| g_s(t+1) < 0.9 \times g_s(t)\right). \quad (1)$$

A gene g_r is a potential co-regulator of a target gene g_s (denoted $g_r \rightharpoonup g_s$) if and only if it exhibits a ±0.1 × g_r change of expression during a change of expression of g_s of ±0.1 × g_s:

$$g_r \rightharpoonup g_s \leftrightarrow (g_r(t) > 1.1{\times}g_r(t) | g_r(t) < 0.9{\times}g_r(t)) \,\&\, \left(g_s(t) > 1.1 \times g_s(t) \middle| g_s(t) < 0.9 \times g_s(t)\right). \quad (2)$$

Step 2.1.2. DBN modeling
We implemented the GRN inference step as a DBN learning problem (2–4). DBNs are BNs where the dependences among the variables can be derived over adjacent time steps. We adopted the common first order Markov assumption and stationarity assumption for deriving the DBN: the value of each variable at one time point depends only on the values of its regulators at the previous time point; and the probability of each variable in relation to its regulators is independent of time. These assumptions allow us to factorize the joint probability distribution:

$$P(X_1, \ldots, X_m) = \prod_i P(X_i | X_1, \ldots, X_{i-1}) = \prod_i P(X_i | \boldsymbol{Pa}(X_i)) \quad (3)$$

where X_i is the expression of gene i, $m = n(T-1)$ is the number of genes (nodes), and $\boldsymbol{Pa}(X_i)$ is the set of regulators of gene i, more generally known as parents of node i.

Given some observations of the variables over time, estimating a DBN consists of finding the structure of (3) that maximizes a score, i.e., finding the most likely parents of each node. The score that we maximize in GENIST is the Bayesian Dirichlet equivalence uniform (BDeu)(5), which is one of the most commonly used scores to evaluate the structure of a DBN given no prior

knowledge. The BDeu score of a DBN can be decomposed as the sum of the scores of the log conditional probabilities of each node. Consequently, we worked with the log of the BDeu, BDeu$_l$:

$$BDeu_l(D,G) = log(P(G)) + \sum_{i=1}^{n}\sum_{j=1}^{q_i}(log(\frac{\Gamma(\frac{\alpha}{q_i})}{\Gamma(\sum_{k=1}^{r_i} N_{ijk}+\frac{\alpha}{q_i})}) + \sum_{k=1}^{r_i} log(\frac{\Gamma(N_{ijk}+\frac{\alpha}{q_i})}{\Gamma(\frac{\alpha}{r_i q_i})}))$$

(4)

where G refers to the Bayesian graph, D refers to the dataset containing the observations of the system, N_{ijk} indicates the number of data vectors in which gene i,X_i, has the value k while its parents are in the jth configuration, and α refers to the hyperparameters of the Dirichlet distribution, which are user-specified.

From (3) and (4) it follows that we can decompose the problem of deriving the DBN into finding the parents for each node. For this, we started by discretizing the expression levels of each gene. The default number of discretization levels is 2, which was heuristically determined to yield the highest prediction precision (the number of discretization levels however, can be user-specified). Then, for each gene, a list of all possible subsets of potential regulators is generated. The maximum size allowed for any subset (maximum number of regulators of a gene) is 3, although it can be user-specified. Therefore, if gene i has r potential regulators and the maximum number of regulators is r_{max}, then the number of combinations of potential regulators, l, equals the number of subsets in r of size $k \leq r_{max}$:

$$l = \sum_{k=0}^{r_{max}}\binom{r}{k} \qquad (5)$$

The $BDeu_l$ is then used l times to evaluate the likelihood that each gene is due to each of the subset of potential regulators. The regulators of gene i are the ones contained in the subset that led to the highest value of the$BDeu_l$.

Step 2.2. Inferring inter-cluster connections
Definition1. *Cluster (sub-network) hub:* Cluster node with the largest degree of edges leaving the node (out-degree).

Input: Time-series gene expression data for all hubs in all clusters C_n, for $n \in [1,c]$.
Output: Sub-network for all the hubs

Repeat *Steps 2.1.1-2.1.2* with time-series gene expression data for all hubs in all clusters, which leads to inter-cluster interactions inferred among the cluster hubs.

Step 2.3 Merging sub-networks
Input: c+1 non-connected sub-networks
Output: GRN formed by c+1 interconnected sub-networks.

Merge the *c+1* sub-networks inferred, corresponding to clusters C_n, for $n \in [1,c]$ and the sub-network inferred for the cluster hubs.

Step 2.4. Determining the sign of the interactions
We implemented a score to estimate whether the inferred interactions are activations or repressions. The score, which is calculated for each edge individually, determines the conditional probability that a gene is active (inactive) given that a parent was active (inactive) in the prior time point, relative to the probability that a gene is active (inactive) given that a parent was inactive (active) in the prior time point. If the first conditional probability is larger (smaller) than the second one, then the parent is accounted as an

activator (repressor). If both probabilities are equal, then the sign of the interaction is accounted as undetermined.

Step 2.5. Filtering out weak interactions
GENIST incorporates a filter to remove the weakest edges. The threshold for removal can vary between 0 and the largest weight of the resulting network and it can be user-specified. The default value of the threshold is 0, i.e., no edges are removed.

Network representation
GENIST returns a MATLAB plot of the inferred GRN. In addition, it returns an output file containing a table with the inferred interactions in a Cytoscape format (6). This table can be imported into Cystoscape for visual representation purposes. The networks presented in this manuscript were plotted via Cytoscape.

Inference of the stem cell GRNs with GENIST
The networks of XYL, QC, and SCN enriched genes presented in this manuscript were inferred by applying GENIST to a spatiotemporal dataset. Specifically, a spatial dataset (QC, CEI (7), XYL, SCN (8)) in combination with transcriptional profiles corresponding to the elongation (Stage II) and differentiation (Stage III) zones of the root (9) were used by the first step of the algorithm (clustering). Then, either the dataset of the transcriptional profile of 12 developmental zones of the Arabidopsis root, which contain temporal information (10), or our stem cell time course dataset, or the shuffled 12 developmental time points were used by the second step of the algorithm (DBN inference).

SI Results

GENIST validation
We tested GENIST's inference step with *in silico* time-series datasets and compared its performance with several other GRN inference algorithms that are designed to work with temporal data. For this, we used the DREAM 4 challenge 2 (11–13). To recover only the most probable interactions, we sought networks with high True Positives/ False Positives (TP/FP) ratios. We chose two performance metrics to indirectly measure this ratio, precision and Area under the Precision Recall Curve (AUPRC). Using these metrics we could score GENIST and compare it to other currently used algorithms (ebdbnet (14), ScanBMA (15), ARACNE (16), CLR (17), MRNET (18), LASSO (19,20)). GENIST performed better than these known algorithms in terms of the precision for both the 10- and 100-node networks (Table S2 (15) & S3 (15)). In the 10-node network, GENIST performed slightly below average in terms of the AUPRC, but it achieved the best performance in terms of AUPRC for 100-node networks. Overall, GENIST consistently achieved the lowest number of false negatives (FN) while maintaining the highest TP/FP ratio across network sizes. The high AUPRC and high precision of GENIST thus, suggested that we could infer networks with a large number of TPs to be validated with biological experiments.

To understand if the application of GENIST to root data would allow us to recover root GRNs, we first applied GENIST to two different root datasets to infer known networks. Specifically, we used the transcriptional profile of 12 developmental time points of the root (10) and a stem cell time course that we generated, both of which embed temporal information about transcriptional regulations of genes expressed in the root. We used root 1 of the 12 developmental time points and replicate 1 of the stem cell time course, as these replicates led to the highest precisions for inferring the known networks. We inferred a phloem (Fig. S3), a CEI (Fig. S4A-F), and a XYL network (Fig. S4G-L), with 5, 4, and 3 nodes, respectively. We illustrated the inferred importance of genes by depicting different node sizes as a function of their outdegree. The use of small networks allowed us to test the Bayesian inference step of GENIST without the need of a clustering step. First, we predicted the phloem network consisting of *ALTERED PHLOEM DEVELOPMENT* (*APL)* and four downstream genes (Fig. S3A)(21). We found that the precision of the phloem network obtained with the 12 developmental time points (Precision = 0.8) (Fig. S3C) was higher than that obtained with the stem cell time course (Precision = 0.45) (Fig. S3D). Similarly, the precision of the XYL network was higher when obtained with the 12 developmental time points

(Precision = 1) (Fig. S4B) than with the stem cell time course (Precision = 0.25) (Fig. S4C). The precision of the CEI network, whose genes are expressed in the stem cells, was similar with both datasets (Precision = 0.8 and 0.71, with the 12 developmental time points and the stem cell time course datasets, respectively) (Fig. S4H,I). To understand if our predicted regulations could be due to random inferences, we shuffled the rows and columns (corresponding to the genes and time zones) of the 12 developmental time points dataset and inferred the same three networks. We obtained precisions of 0.33, 0.5, and 0 for the phloem, CEI and XYL networks, respectively. This indicated that although some inferences could be correctly guessed, the 12 developmental time points and the stem cell time course network precisions are higher than those inferred with random data. In addition, our results indicated that the 12 developmental time points dataset led to higher precision networks and it can be used to infer regulations of genes during both stem cell processes and later root development. Next, we compared the performance of GENIST with two other well-known and most often used algorithms, ARACNE and CLR (22) (Fig. S3F,G). We first predicted the phloem network (Fig. S3A) (21) (Fig. S3C). Since, overall, using the 12 developmental time points performed better than using the stem cell time course for inferring the small networks, we used the 12 developmental time points of these five genes (10) as the input data into the three algorithms (Fig. S3B). Our inferred network (Precision = 0.8, Fig. S3C) outperformed the other algorithms (Precision = 0.2, Precision = 0, Fig. S3F,G). Similarly, GENIST outperformed the other algorithms in the 4 and 3 node networks (Fig. S4). Thus, these results validate the performance of our algorithm in identifying small GRNs and confirm that the 12 developmental time points dataset can be used in combination with GENIST to infer root GRNs.

Given the expected complexity of the stem cell enriched gene networks and the anticipation that they would contain genes involved in multiple cells, we tested whether GENIST could infer relationships among genes expressed in specific stem cells or across more than one stem cell. For this, we combined the previously tested phloem, CEI, and XYL networks to test GENIST. As genes expressed in a specific cell type most likely interact with genes expressed in the same cell, we hypothesized that a preprocessing step prior the Bayesian inference, where genes were grouped based on their co-expression, could be biologically relevant. We, therefore, incorporated into the algorithm a clustering step. We used the full GENIST algorithm, with both clustering and inference steps, to infer relationships among the combination of the phloem, CEI, and XYL networks (Fig. S4M). For this, we used our cell type specific dataset to group all the genes from these networks according to their co-expression across the stem cells and Stage II and Stage III. GENIST grouped these cell-type-specific genes in 3 clusters and used the 12 developmental time points to build one Bayesian Network for each resulting cluster of genes. This step resulted in three independent Bayesian Networks (cluster partial precision = {0.75, 0.8, 0.5}). GENIST then predicted relationships among the most representative nodes (hubs) of each cluster to infer inter-cluster edges (SI Methods). Therefore, inter-cluster edges are representative of potential regulations among co-expressed genes. As a result, the hubs in our networks have a regulatory effect on both their own cell-type-specific sub-network and the entire network. We identified the hubs of each sub-network as *SHR (XYL)*, *SCR (CEI)*, and *APL (PHL)*. Inferring the relationships between *SHR, SCR,* and APL connected the clusters and resulted in the final network (Fig. S3H) (Precision = 0.5556), which improved the results from applying GENIST's inference step alone (Precision = 0.071). With an overall ~8-fold precision improvement, our approach suggests that incorporating a clustering step using a cell-specific dataset can exceed the performance of temporal data based-inference algorithms. Moreover, the clustering step refined the computational complexity of the network inference. In particular, the time complexity as the function of the average number of genes in a cluster increased at a rate of ~ $O(n^2\log^3(n))$, corresponding to a 2627-fold change between no clustering and 15 clusters for a 100-node network (Table S4). Thus, clustering computationally facilitates the application of GENIST to large networks. Overall, our result indicated that the capacity of GENIST to integrate distinct datasets can be beneficial for inferring GRNs in organisms where transcriptional datasets of diverse characteristics are available, and in particular, for inferring GRNs of genes enriched in the stem cells.

Mathematical modeling

To predict the sign of the undetermined regulations in the *PAN* subnetwork inferred with the 12 developmental time points of the root (Fig. 3A), and in turn, provide a system that permitted to simulate the predictions made by our network, we developed a mathematical model that represents all the

dependences inferred between *PAN* and its downstream targets (Fig. 3A). The model development consisted of a first step to determine the unknown regulation signs (finding the optimal configuration of the model) and a second step to find the parameters of the optimal configuration of the model.

Our model is comprised of eight Ordinary Differential Equations (ODEs), each of which represents the change of one TF expression over time. In our ODEs, we assumed independent probabilities of regulatory bindings to describe the equations where TFs are influenced by multiple factors. Every dependency of a TF on another TF is represented by a Hill function term (representing an activation or a repression), as Hill functions provide a simple platform that captures important aspects of transcriptional regulatory interactions. However, some of the dependences in these equations have an undetermined sign, and the Hill term for these dependences can take two different forms. Specifically, since the *PAN* subnetwork comprises 11 unknown sign regulations, our model can take 2048 configurations.

To understand which configuration would better fit the experimental data, we estimated our model parameters for multiple configurations by fitting the ODEs to the gene expression patterns along the 12 developmental time points of the root. We particularly used a global optimization method, Simulated Annealing, to fit the ODEs to the 12 developmental time points of the root. Since performing an exhaustive search throughout 2048 configurations would be unfeasible, we took a sub-optimal approach by fitting one equation at a time. In particular, during the search of the possible configurations of ODE i involving n unknown regulations, we fitted the parameters of the equation for the 2^n possible configurations, while we eliminated the dependences among the rest of the ODEs. To eliminate the dependences while fitting the parameters of ODE i, we used a first order approximation of the experimental data on all the factors associated to all other ODEs. We found that this provided more robust optimization results than optimizing all equations at once. For ODE i, we selected the configuration that minimized the mean squared error between the experimental and simulated expression value of factor i.

To ensure that all parameters were jointly estimated, we performed a final parameter estimation of the model by fitting all the ODEs to the 12 developmental time points of the root. The resulting model and associated parameters are described in (6) – (13) and Table S6. This model allowed us to perform simulations of wild type and *pan* mutant backgrounds (Fig. 3 B-C). These simulations permitted the comparison of our results to experimental validations to measure the confidence of the network predictions.

$$\boldsymbol{PAN}: \quad \frac{dP}{dt} = p_P - d_P P \tag{6}$$

$$\boldsymbol{BRAVO}: \frac{dB}{dt} = p_B \frac{P^{b_1}}{K_{B_1}^{b_1} + P^{b_1}} \frac{N^{b_2}}{K_{B_2}^{b_2} + N^{b_2}} \frac{WP^{b_3}}{K_{B_3}^{b_3} + WP^{b_3}} \frac{E^{b_4}}{K_{B_4}^{b_4} + E^{b_4}} \frac{HC^{b_5}}{K_{B_5}^{b_5} + HC^{b_5}} - d_B B \tag{7}$$

$$\boldsymbol{NTT}: \quad \frac{dN}{dt} = p_N \frac{P^{n_1}}{K_{N_1}^{n_1} + P^{n_1}} \frac{B^{n_2}}{K_{N_2}^{n_1} + B^{n_2}} \frac{E^{n_3}}{K_{N_3}^{n_3} + E^{n_3}} - d_N N \tag{8}$$

$$\boldsymbol{WRKY23}: \frac{dW}{dt} = p_W \frac{B^{w_1}}{K_{W_1}^{w_1} + B^{w_1}} \frac{P^{w_2}}{K_{W_2}^{w_2} + P^{w_2}} \frac{1}{1 + (\frac{N}{K_{W_3}})^{w_3}} - d_W W \tag{9}$$

$$\boldsymbol{HSFA1E}: \frac{dHA}{dt} = p_{HA} \frac{B^{ha_1}}{K_{HA_1}^{w_1} + B^{ha_1}} \frac{P^{ha_2}}{K_{HA_2}^{ha_2} + P^{ha_2}} \frac{1}{1 + (\frac{N}{K_{HA_3}})^{ha_3}} - d_{HA} HA \tag{10}$$

$$\boldsymbol{WIP4}: \frac{dWP}{dt} = p_{WP} \frac{1}{1 + (\frac{N}{K_{WP_1}})^{wp_1}} \frac{E^{wp_2}}{K_{WP_2}^{wp_2} + E^{wp_2}} - d_{WP} WP \tag{11}$$

$$\boldsymbol{HSFC1}: \frac{dHC}{dt} = p_{HC} \frac{B^{hc_1}}{K_{HC_1}^{hc_1} + B^{hc_1}} \frac{N^{hc_2}}{K_{HC_2}^{hc_2} + N^{hc_2}} - d_{HC} HC \quad \textbf{(12)}$$

$$\boldsymbol{EIN3}: \frac{dE}{dt} = p_E \frac{HC^{e_1}}{K_{E_1}^{e_1} + HC^{e_1}} \frac{1}{1+(\frac{N}{K_{E_2}})^{e_2}} \frac{1}{1+(\frac{B}{K_{E_3}})^{e_3}} - d_E E \quad \textbf{(13)}$$

To understand if the stem cell time course and the 12 developmental time points networks would lead to similar conclusions, we next modelled the *PAN* subnetwork inferred with the stem cell time course (Fig. S6B). We developed the model as explained for the one derived from the 12 developmental time points. The resulting 9-ODE model and associated parameters are described in (14) – (22) and Table S7. This model allowed us to perform simulations of wild type and *pan* mutant backgrounds (Fig. S6 C-D).

$$\boldsymbol{PAN}: \quad \frac{dP}{dt} = p_P - d_P P \quad \textbf{(14)}$$

$$\boldsymbol{BRAVO}: \frac{dB}{dt} = p_B \frac{P^{b_1}}{K_{B_1}^{b_1} + P^{b_1}} \frac{1}{1+(\frac{EB}{K_{B_2}})^{b_2}} \frac{1}{1+(\frac{E}{K_{B_3}})^{b_3}} - d_B B \quad \textbf{(15)}$$

$$\boldsymbol{NTT}: \quad \frac{dN}{dt} = p_N \frac{P^{n_1}}{K_{N_1}^{n_1} + P^{n_1}} \frac{HC^{n_2}}{K_{N_2}^{n_1} + HC^{n_2}} \frac{1}{1+(\frac{W}{K_{N_3}})^{n_3}} - d_N N \quad \textbf{(16)}$$

$$\boldsymbol{WRKY23}: \frac{dW}{dt} = p_W \frac{S^{w_1}}{K_{W_1}^{w_1} + S^{w_1}} \frac{1}{1+(\frac{E}{K_{W_2}})^{w_2}} - d_W W \quad \textbf{(17)}$$

$$\boldsymbol{HSFA1E}: \frac{dHA}{dt} = p_{HA} \frac{B^{ha_1}}{K_{HA_1}^{w_1} + B^{ha_1}} \frac{1}{1+(\frac{EB}{K_{HA_2}})^{ha_2}} - d_{HA} HA \quad \textbf{(18)}$$

$$\boldsymbol{SCRM2}: \frac{dS}{dt} = p_S \frac{1}{1+(\frac{E}{K_{S_1}})^{s_1}} \frac{W^{s_2}}{K_{S_2}^{s_2} + W^{s_2}} - d_S S \quad \textbf{(19)}$$

$$\boldsymbol{HSFC1}: \frac{dHC}{dt} = p_{HC} \frac{B^{hc_1}}{K_{HC_1}^{hc_1} + B^{hc_1}} \frac{N^{hc_2}}{K_{HC_2}^{hc_2} + N^{hc_2}} - d_{HC} HC \quad \textbf{(20)}$$

$$\boldsymbol{EIN3}: \frac{dE}{dt} = p_E \frac{1}{1+(\frac{P}{K_{E_1}})^{e_1}} \frac{1}{1+(\frac{N}{K_{E_2}})^{e_2}} \frac{1}{1+(\frac{HC}{K_{E_3}})^{e_3}} - d_E E \quad \textbf{(21)}$$

$$\boldsymbol{EBP}: \frac{dEB}{dt} = p_{EB} \frac{1}{1+(\frac{P}{K_{EB_1}})^{eb_1}} \frac{1}{1+(\frac{B}{K_{EB_2}})^{eb_2}} \frac{E^{eb_3}}{K_{EB_3}^{eb_3} + E^{eb_3}} - d_{EB} EB \quad \textbf{(22)}$$

Topological characteristics of the network of genes enriched in the stem cells

An important feature of complex biological networks are developmental motifs such as positive and negative feedback loops (FLs) (23,24), which are more frequent in transcriptional regulatory networks than in random networks (25,26). FLs are known to be involved in multiple regulatory processes in multicellular organisms, where they tend to establish oscillations and bistability. Therefore, we analyzed whether our networks would recapitulate a significant number of FLs that could validate that our network did not have the characteristics of random networks. To test this, we simulated 1000 random networks of the size of our network. This randomization resulted in an average of 8.5 FLs/network, while our network contains a total of 151 FLs. This indicates that the number of FLs inferred in our network is significantly

higher than expected by random, supporting that bistability or oscillations are behaviors that could emerge more frequently than in random systems and differentiating our network characteristics from those of random networks.

Another common feature of biological networks is that they are scale-free, meaning that the distribution of the number of edges leaving each node (outdegree distribution) fits a decaying power law (27). This is in contrast to random networks, where the outdegree distribution is Gaussian (27). We found that our network showed the characteristics of a scale-free network (Fig. S9B), which indicates that there is a small number of genes with extremely high outdegree. Genes that have high outdegree in cellular networks tend to have an important biological function (28). Thus, the absence of these genes with high outdegree (hubs) could lead to functional or morphological disruptions of the stem cells. This characteristic suggests that our network topology can be used to guide the experimental search for stem cell regulators. Accordingly, we found that some of the predicted hubs have important roles in regulating the root SCN and meristem development, such as SHR (7,29–34), BBM(35), and AIL6 (PLT3)(35), as well as PAN (Fig. S6) (this study).

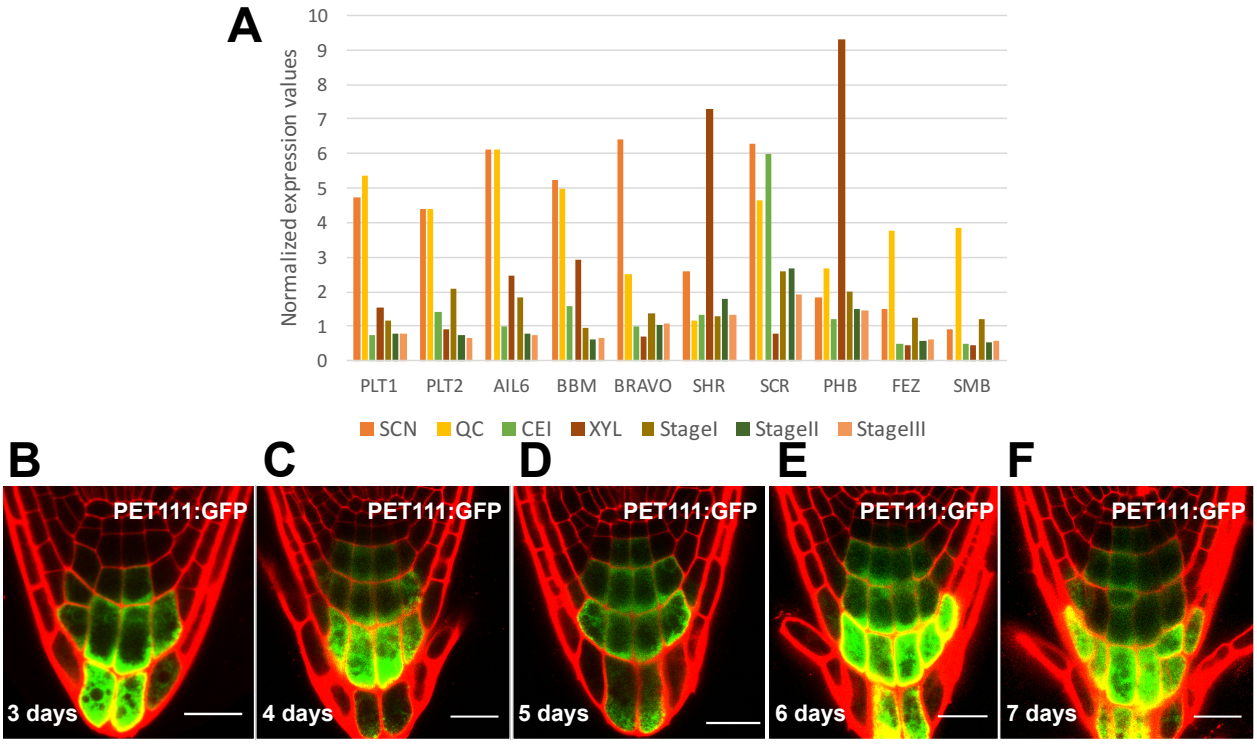

Fig. S1. A ANOVA normalized expression values of known stem cell factors in SCN, QC, CEI, XYL Stage I to III. **B-F** Marker line (PET111:GFP) (10) used for the stem cell time course at days 3, 4, 5, 6, and 7, where the GFP negative cells were sorted. Scale bars are 20 μm.

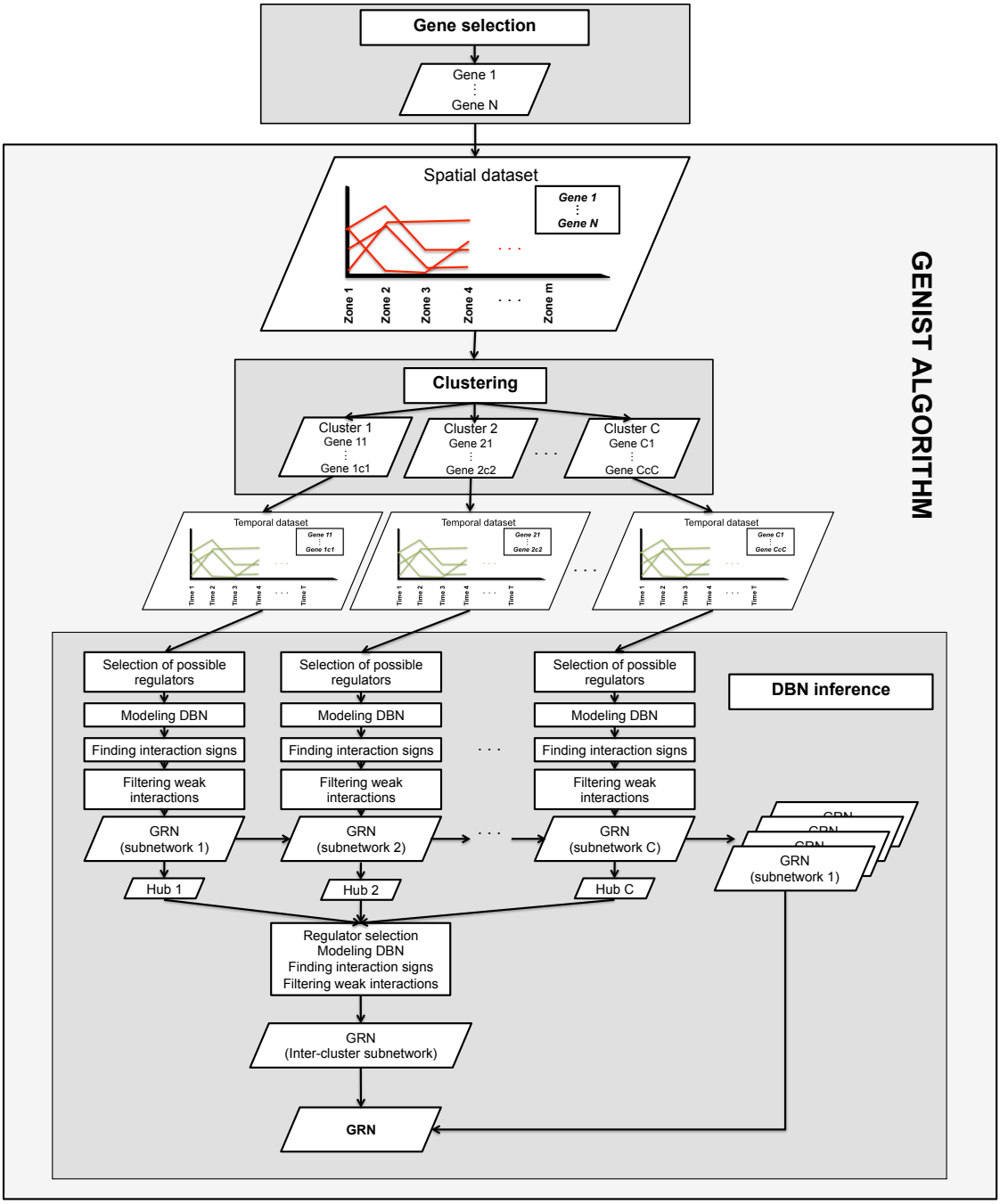

Fig. S2. GENIST block diagram. GENIST is implemented in MATLAB, and is composed of two consecutive steps, clustering and GRN inference. Clustering is performed based on a spatial dataset. Each resulting cluster is independently processed by the GRN inference step, based on a temporal dataset.

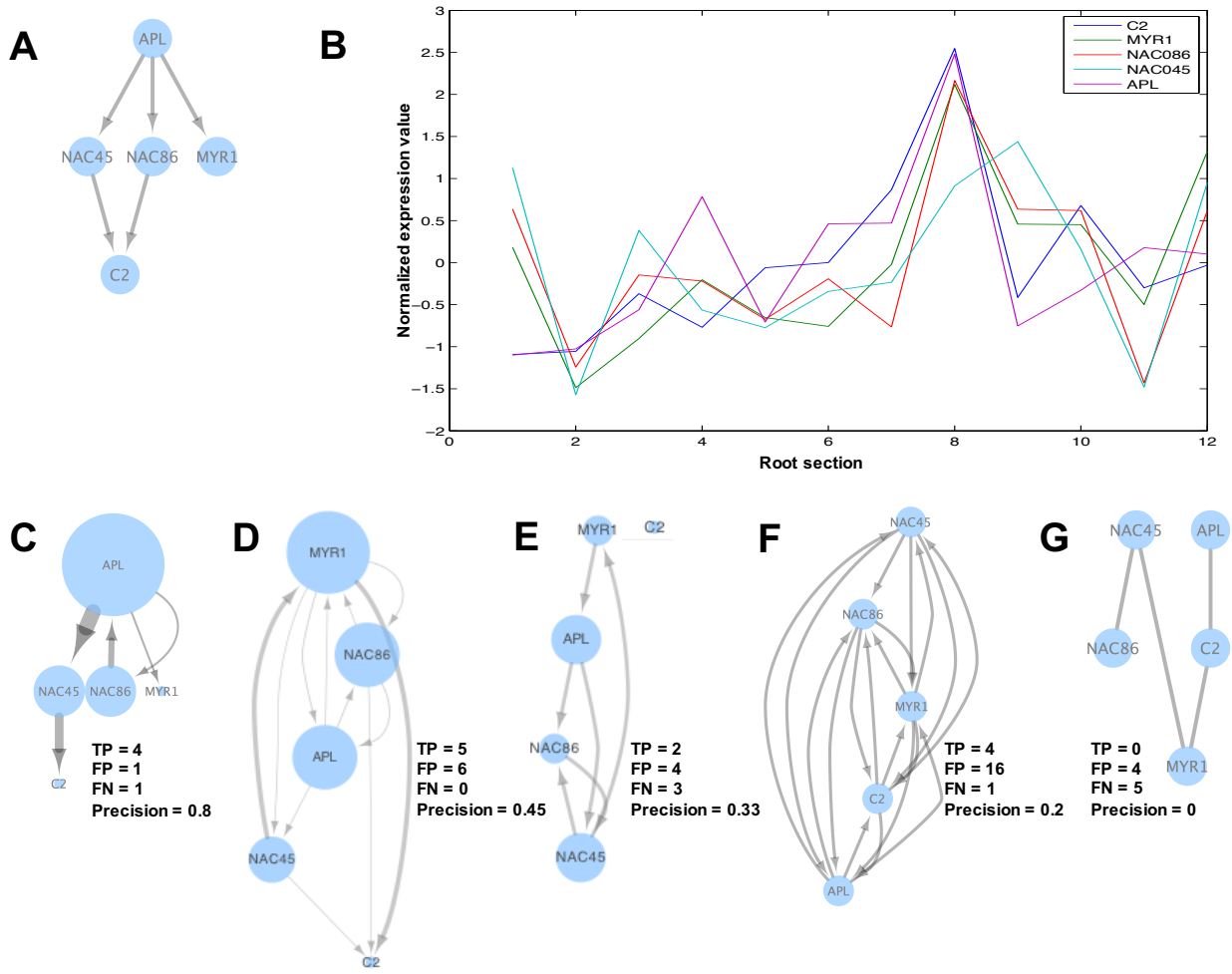

Fig. S3. A-D Validation of GENIST's inference step through the prediction of a small known Arabidopsis network composed of *APL* and its downstream targets *NAC45, NAC86, MYR1*, and *C2*, and comparison of its performance with some homologous algorithms. **A** Known genetic interactions among *APL, NAC45, NAC86, MYR1,* and *C2*. **B** Input data into the inference algorithms, consisting of the expression levels across 12 developmental time points (10) of the five TFs. **C-E** Networks inferred by GENIST's inference step. The size of the nodes correlates with the amount of direct targets, illustrating the regulatory capacity of the node. The thickness of the edges illustrates the confidence of the regulation. **C** Network inferred with the 12 developmental time points. **D** Network inferred with the stem cell time course. **E** Network inferred with the shuffled 12 developmental time points. **F** CLR's inferred network. **G** GRN inferred by ARACNE. **C-G** The number of TP, FP, and FN inferred edges shown, as well as precision of the inferred network, are in relation to the known genetic interactions.

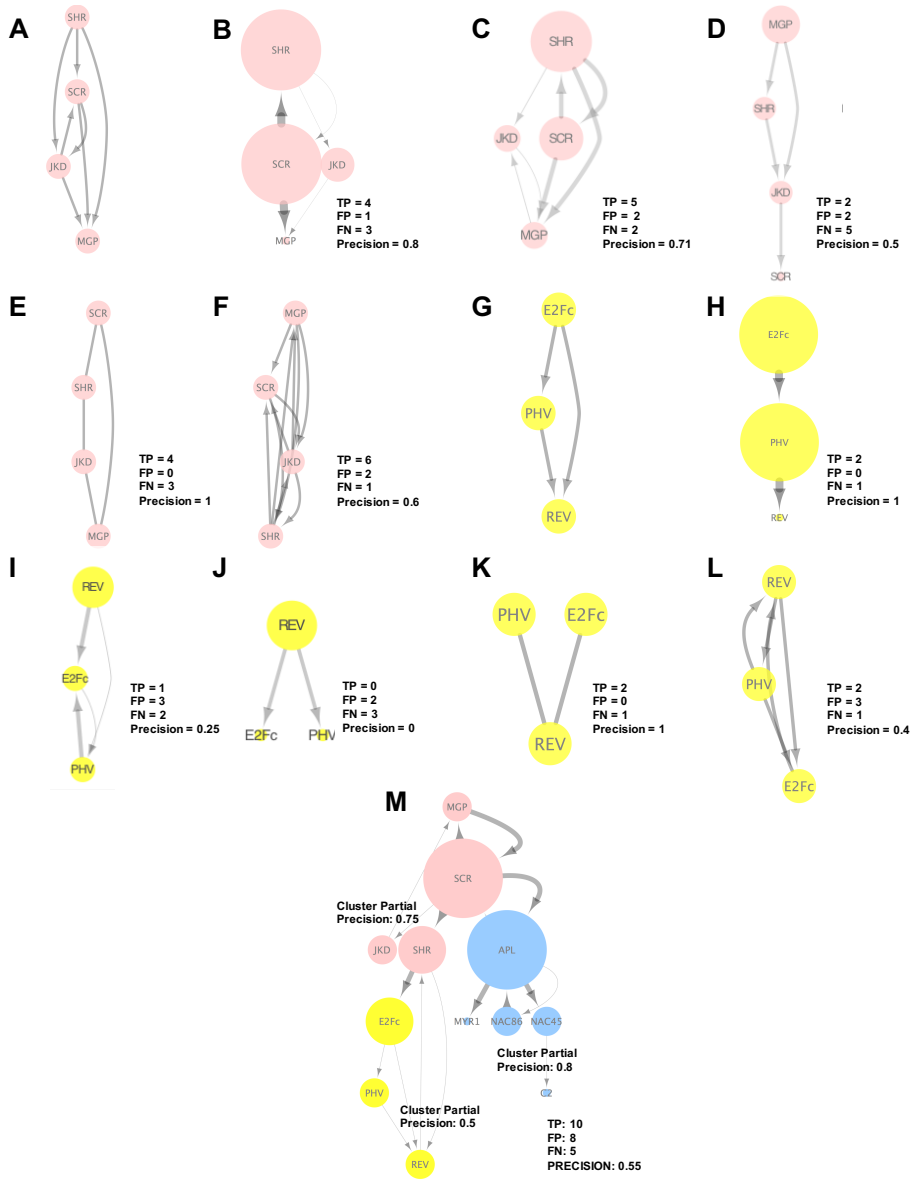

Fig. S4. Validation of GENIST using two small known Arabidopsis networks. **A-F** Prediction of a small known Arabidopsis network composed of *SHR, SCR, JKD,* and *MGP*, and comparison of its performance with some homologous algorithms **A** Known genetic interactions among *SHR, SCR, JKD,* and *MGP*. **B** Network inferred by GENIST's inference step with the 12 developmental time points dataset . **C** Network inferred by GENIST's inference step with the stem cell time course. **D** Network inferred by GENIST's inference step with the shuffled 12 zones of the longitudinal axis of the Arabidopsis root. **E** Skeleton inferred by ARACNE. **F** CLR's inferred network. **G-L** Prediction of a small known Arabidopsis network composed of *E2Fc, PHV*, and *REV*, and comparison of its performance with some homologous algorithms. **G** Known genetic interactions among *E2Fc, PHV*, and *REV*. **H** Network inferred by GENIST's inference step with the 12 developmental time points dataset . **I** Network inferred by GENIST's inference step with the stem cell time course. **J** Network inferred by GENIST's inference step with the shuffled 12 developmental time points dataset . **K** Skeleton inferred by ARACNE. **L** CLR's inferred network. **M** Evaluation of GENIST performance on a 12-node network. The network contains three independent sub-networks, corresponding to the networks shown in Fig. S2 and S3. The clustering step, performed based on our cell-specific dataset, allows the algorithm to initially divide the network into smaller sub-networks, reducing the number of potential regulators and leading to a precision of 0.55. Although the clustering step fails to group SHR with MGP, SCR, and JKD, the link between cluster hubs connects it to the correct cluster. **A-M** The number of TP, FP, and FN inferred edges shown, as well as precision of the inferred networks, are in relation to the known genetic interactions. **B-D,H-J** The size of the nodes correlates with the amount of direct targets, illustrating the regulatory capacity of the node. The thickness of the edges illustrates the confidence of the regulation.

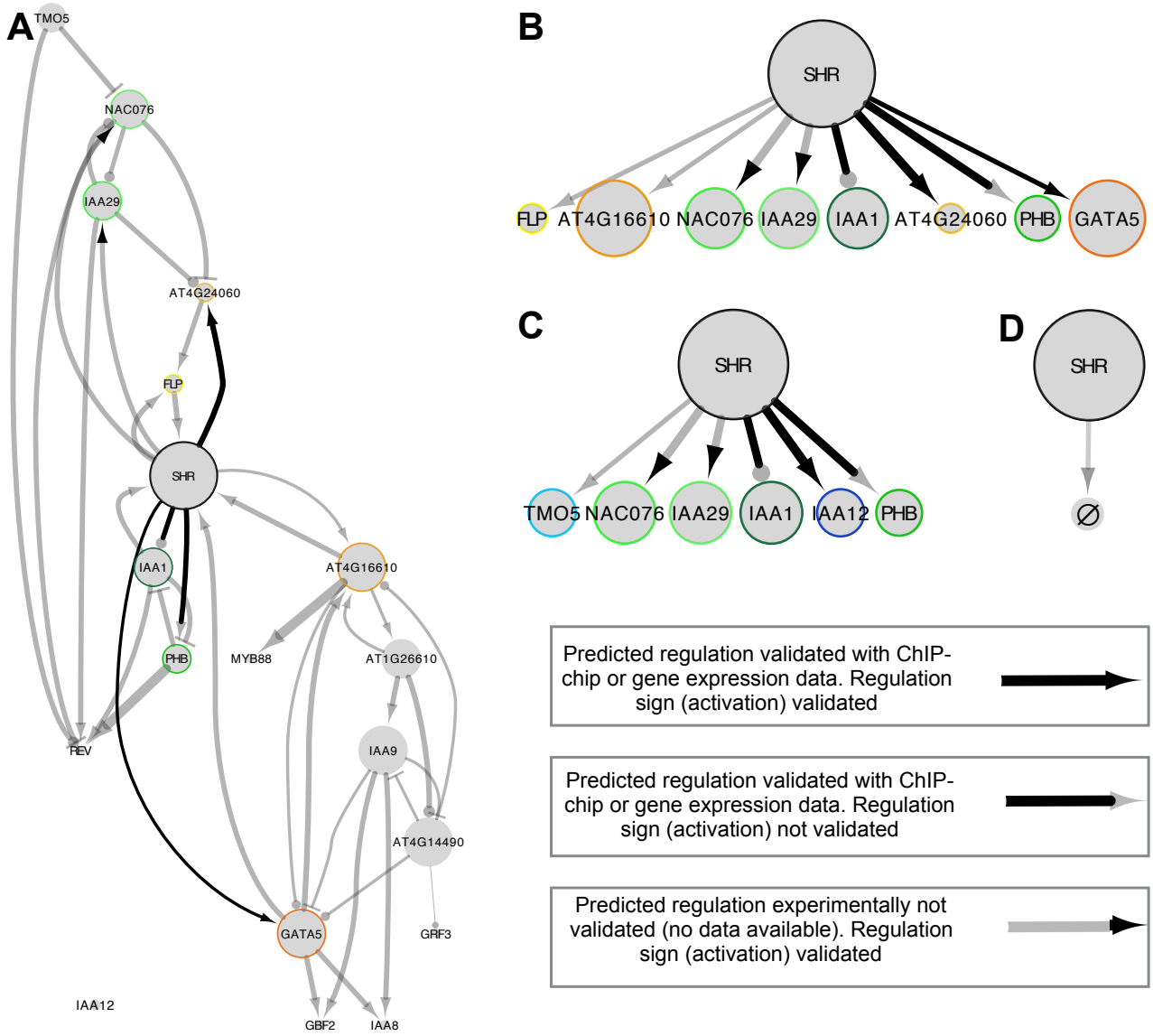

Fig. S5. Network of XYL-enriched TFs. **A** Network inferred with the 12 developmental time points dataset. Node sizes indicate importance of the nodes in terms of the number of TFs that they regulate. Nodes are color-coded. Green tones represent SHR targets inferred in common between the 12 developmental time points and the stem cell time course datasets. Orange tones represent SHR targets uniquely inferred with the 12 developmental time points dataset. Blue tones represent SHR targets uniquely inferred with the stem cell time course dataset. **B-D** Subset of the XYL network showing *SHR* and its predicted targets **B** Network inferred with the 12 developmental time points dataset . **C** Network inferred with the stem cell time course. Four of the targets (NAC076, IAA29, IAA1, PHB) are in common between B and C. **D** Network inferred with the shuffled 12 developmental time points dataset . No targets are predicted for *SHR*.

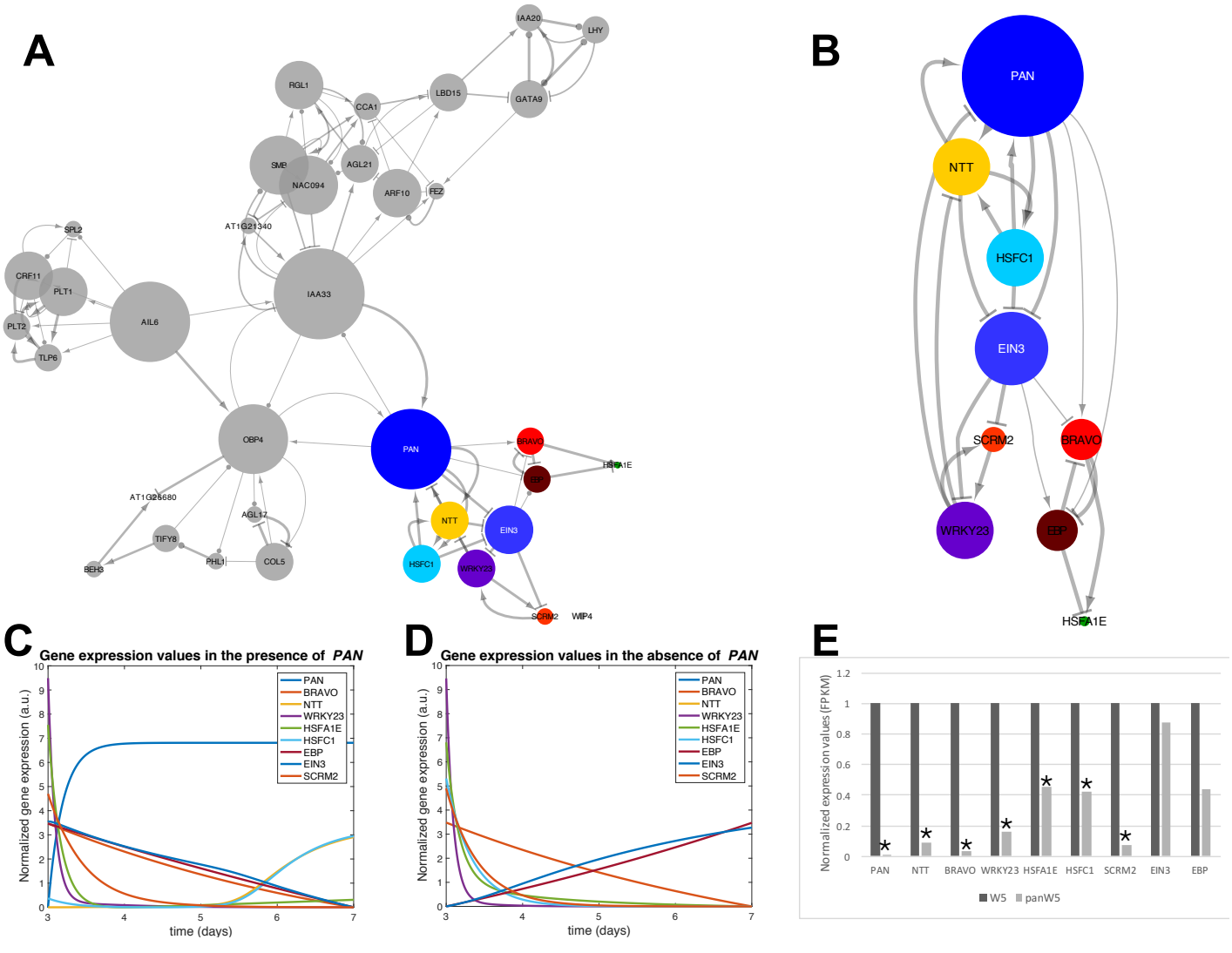

Fig. S6. Network of QC-enriched TFs. **A** QC network inferred with the stem cell time course. **B** Optimal configuration (combination of sings –activation or repression– of the regulations that were inferred with undefined signs, which best fits the data in the simulations of the equations) of the subnetwork of *PAN* and its downstream targets. **C-D** Resulting expression values of *PAN* and its downstream targets, over time (days 3 to 7), after simulating the optimal configuration of the model. The expression of each gene through time (in FPKM) was normalized (0 mean 1 variance and shifted into the positive quadrant) for plotting purposes, as FPKM values of different genes can have different scales. **C** Model simulated with the fitted equation parameters. **D** Model simulated with the *PAN* associated parameters set to zero to simulate a *pan* mutant situation. **E** Normalized expression values of *PAN* and its predicted downstream targets in Col-0 wild-type and in a *pan* mutant. Star (*) represents statistically significant changes of expression between the mutant and the wild type ($q < 0.05$).

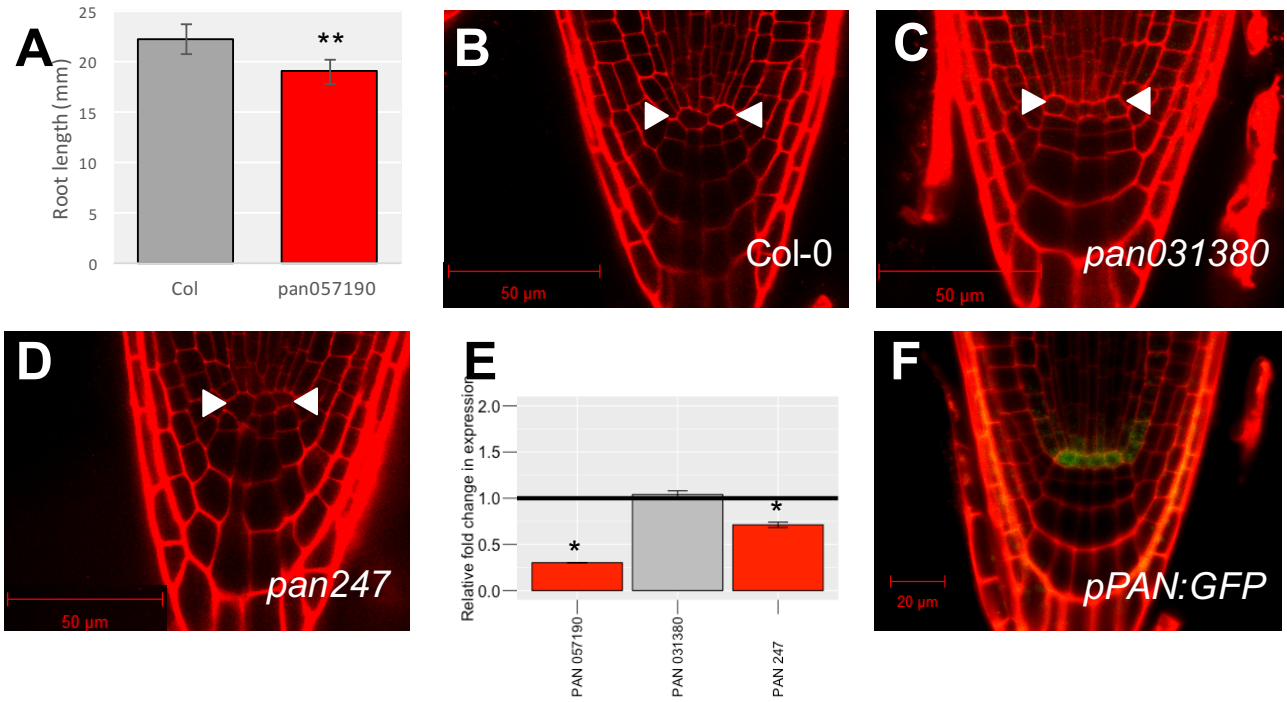

Fig. S7. *pan* mutant alleles. **A** Root length (mean ± standard deviation) of 7-day-old primary seedling roots of Col-0 (n = 16) and *pan* mutant alleles (n = 16). Double star (**) represents statistical difference ($p < 10^{-4}$, Student t-test). **B-D** *pan* mutant confocal images showing a disorganized QC of two additional alleles. **E** Relative fold-change of expression of *PAN* in the *pan* mutant alleles. Star (*) represents statistical difference ($p < 0.05$, Student t-test). **F** Confocal image showing the transcriptional *pPAN:GFP* fusion.

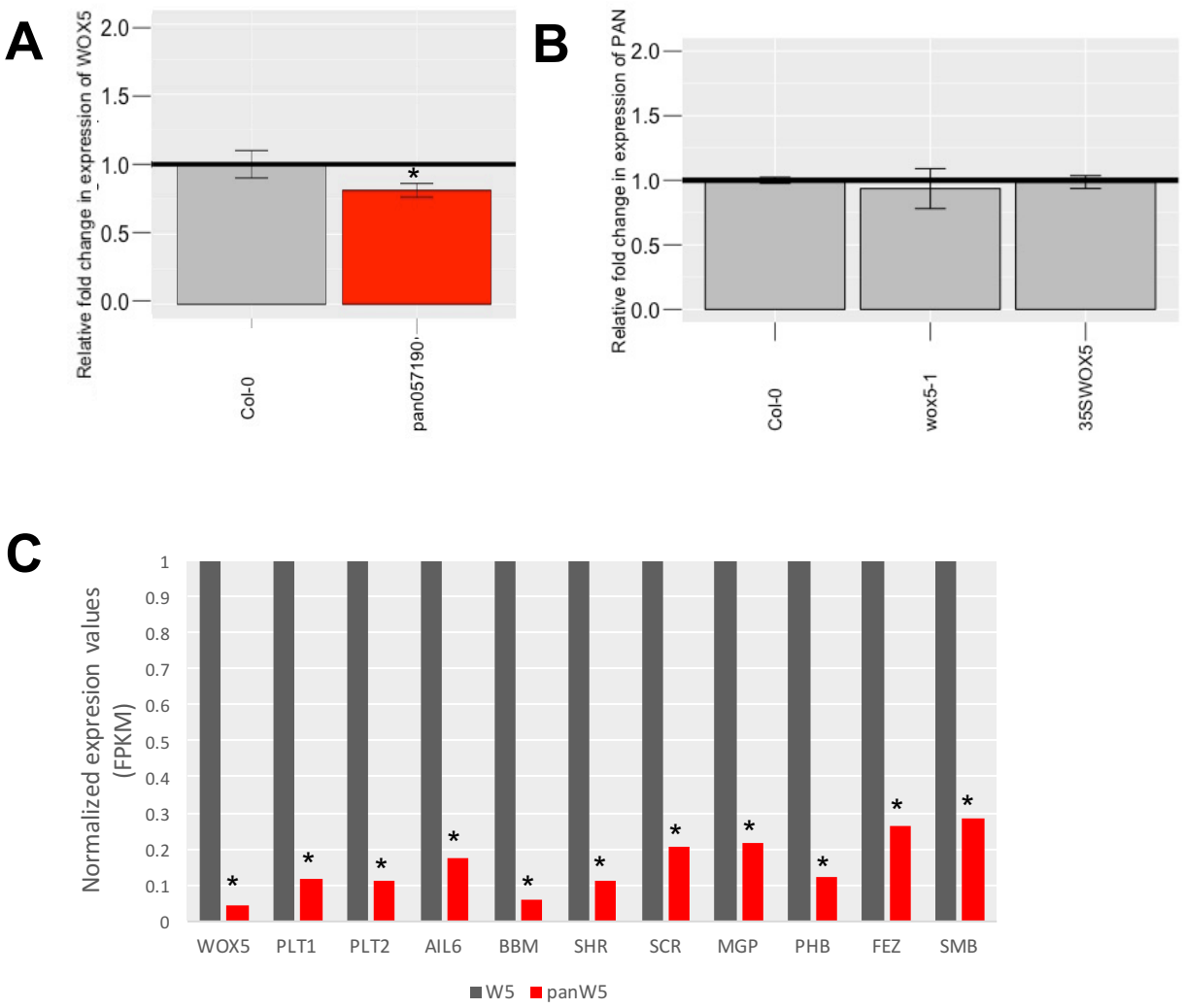

Fig. S8. A qRT-qPCR results of *WOX5* in a *pan* mutant background; star (*) represents significant change of expression of *WOX5* (Student t-test, $p < 0.05$). **B** qRT-qPCR results of *PAN* in a *wox5* mutant and *35S:WOX5* lines showing no significant changes of expression. **C** Normalized expression values of known stem cell regulators in Col-0 wild-type and in a *pan* mutant; star (*) represents statistically significant changes of expression between the mutant and the wild type ($q < 0.05$).

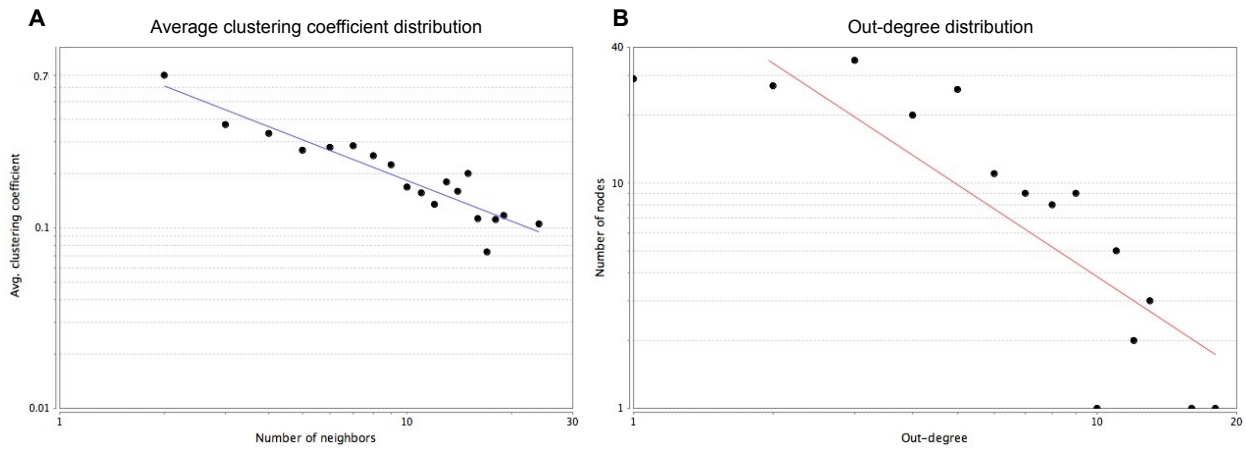

Fig. S9. Topological information of the stem cell maintenance network comprising the 201 TF found to be enriched in the SCN. The solid lines correspond to the power laws that the distributions fit. **A** Average clustering coefficient distribution, which follows a power law of the form $y = 1.027x^{-0.75}$. **B** Out-degree distribution, which follows a power law of the form $y = 86.906x^{-1.355}$.

Table S1. Gene Ontology category enrichment analysis in each of the stem cell types.

GO TERM	CEI	XYL	QC	SCN
GDP-mannose 4,6-dehydratase activity	-3.02	0	0	0
De novo' GDP-L-fucose biosynthesis	-3.32	0	0	0
Biological process unknown	0	-3.18	0	0
Microtubule motor activity	0	-9.29	0	0
cyclin-dependent protein kinase activity	0	-3.39	0	0
Nucleus	0	-3.61	0	-9.55
Nucleolus	0	-3.65	0	0
Spindle	0	-3.97	0	0
Microtubule associated complex	0	-7.30	0	0
Microtubule-based movement	0	-7.58	0	0
Phragmoplast	0	-3.07	0	0
Determination of bilateral symmetry	0	-3.01	0	0
Guard cell differentiation	0	-3.78	0	0
Histone phosphorylation	0	-3.78	0	0
Transcription factor activity	0	0	-5.69	-16.22
Pattern specification	0	0	-7.43	0
Regulation of transcription	0	0	-3.21	0
Cyclin binding	0	0	-3.15	0
Root development	0	0	-5.09	-3.43
Response to heat	0	0	0	-12.64
Chloroplast	0	0	0	-4.94
Response to wounding	0	0	0	-3.48
DNA binding	0	0	0	-8.30
Beta-fructofuranosidase activity	0	0	0	-3.42
Trehalose biosynthesis	0	0	0	-3.11
Regulation of transcription, DNA-dependent	0	0	0	-9.81
Response to ethylene stimulus	0	0	0	-3.25
Abscisic acid mediated signaling	0	0	0	-3.09
Jasmonic acid mediated signaling pathway	0	0	0	-3.64
Ethylene mediated signaling pathway	0	0	0	-3.77
Transcriptional activator activity	0	0	0	-3.14
Transcriptional repressor activity	0	0	0	-3.86
Positive regulation of transcription	0	0	0	-3.46
Defense response to fungus	0	0	0	-3.86

Values are expressed in a log10 scale. The analysis was performed with the Chip enrich software.

Table S2. Performance of the algorithm, applied to DREAM 4 challenge 2, 10 node networks.

Method	Precision	AUPRC	TP	FP
GENIST	0.520	0.413	26	24
ebdbnet	0.509	0.438	28	27
ScanBMA	0.432	0.505	35	46
ARACNE	0.304	0.388	35	80
CLR	0.215	0.397	50	183
MRNET	0.215	0.409	53	193
LASSO	0.190	0.487	62	265

The true positives (TP) and false positives (FP) are given for the 5 networks. There are a total of 71 true edges across the 5 networks.

Table S3. Performance of the algorithm, applied to DREAM 4 challenge 2, 100 node networks.

Method	Precision	AUPRC	TP	FP
GENIST	0.252	0.139	126	374
ScanBMA	0.153	0.101	193	1062
ARACNE	0.114	0.106	208	1621
ebdbnet	0.054	0.043	182	3201
CLR	0.035	0.123	678	18669
LASSO	0.035	0.073	571	15757
MRNET	0.035	0.13	689	18784

The true positive (TP) and false positive (FP) columns are totaled across all 5 networks. There are 1024 true edges across the 5 networks.

Table S4. GENIST running times

DREAM4 100-gene network 5. 10 runs average	
Number of clusters	**Running time (s)**
No clustering	75804.9981
5 clusters	173.3386
10 clusters	33.2263
15 clusters	28.8517

GENIST running times for processing the DREAM 4 challenge 2, 100-node network 5, for different number of clusters. Times shown (in seconds) are for an average of 10 runs.

Table S5. List of primers

Gene	Purpose	Stock ID	Left Primer	Right Primer
PAN_057190	Genotype	Salk_057190	ACATCAACACGGCCAAGTAAC	TCTCTCCTCACTCCCTCCTTC
PAN 031380	Genotype	Salk_031830	CGGTAACACACATGACACATATG	ATGGTGAAAACCATTGACTGG
PAN 247	Genotype	Sail_247 C12	TTGCCTCAATAAATCAGCCTG	GAATTCTTGGCAGACACTTCG
PP2A	qRT-PCR	X	TAACGTGGCCAAAATGATGC	GTTCTCCACAACCGCTTGGT
PAN	qRT-PCR	X	GGCTTGCACAGCTAGAGGAAGAGC	TCCGGCTGCCAAATGCGTGT
WOX5	qRT-PCR	X	GATTGTCAAGAGGAAGAGAAGGTGA	AGCTTAATCGAAGATCTAATGGCG

Table S6. 12 developmental times model parameters

Symbol	Description	Unit	Value	Range
p_P	PAN production rate	a.u. h^{-1}	1.9165	0.5606 - 3.7094
p_B	BRAVO production rate	a.u. h^{-1}	3.3466	1.4586 - 4.4444
p_N	NTT production rate	a.u. h^{-1}	0.921	0.7834 - 3.9527

p_W	WRKY23 production rate	a.u. h^{-1}	4.1898	0.7636 - 4.1898
p_{HA}	HCFA1E production rate	a.u. h^{-1}	3.1358	1.607 - 4.62
p_{HC}	HSFC1 production rate	a.u. h^{-1}	0.6743	0.6743 - 4.5579
p_{WP}	WIP4 production rate	a.u. h^{-1}	1.7706	0.5837 - 4.3632
p_E	EIN3 production rate	a.u. h^{-1}	3.5084	1.38 - 4.4642
d_P	PAN decay rate	h^{-1}	0.2573	0.1891 - 0.4697
d_B	BRAVO decay rate	h^{-1}	0.1912	0.1826 - 0.3083
d_N	NTT decay rate	h^{-1}	0.131	0.131 - 0.4169
d_W	WRKY23 decay rate	h^{-1}	0.1478	0.0147 - 0.2569
d_{HA}	HCFA1E decay rate	h^{-1}	0.191	0.1431 - 0.3742
d_{HC}	HSFC1 decay rate	h^{-1}	0.2	0.1083 - 0.3054
d_{WP}	WIP4 decay rate	h^{-1}	0.1716	0.1119 - 0.3287
d_E	EBP decay rate	h^{-1}	0.1577	0.1388 - 0.2764
K_{B_1}	HC concentration of half-max HC-dependent B activation	a.u.	0.4874	0.4228 - 0.9508
K_{B_2}	P concentration of half-max P-dependent B inhibition	a.u.	0.3774	0.1226 - 1.2123
K_{B_3}	N concentration of half-max N-dependent B inhibition	a.u.	0.9475	0.2755 - 1.2
K_{B_4}	P concentration of half-max P-dependent B inhibition	a.u.	0.765	0.141 - 0.8236
K_{B_5}	N concentration of half-max N-dependent B inhibition	a.u.	0.9305	0.2397 - 1.0971
K_{N_1}	P concentration of half-max P-dependent N inhibition	a.u.	0.8679	0.25 - 1.1278
K_{N_2}	B concentration of half-max B-dependent N inhibition	a.u.	0.9656	0.1722 - 1.1285
K_{N_3}	E concentration of half-max E-dependent N	a.u.	0.6934	0.2073 - 0.9723

	inhibition			
K_{W_1}	N concentration of half-max N-dependent W activation	a.u.	0.7876	0.3948 - 1.0207
K_{W_2}	P concentration of half-max P-dependent W inhibition	a.u.	0.6272	0.2074 - 0.7628
K_{W_3}	E concentration of half-max E-dependent W activation	a.u.	0.0957	0.0266 - 0.8586
K_{HA_1}	N concentration of half-max N-dependent HA activation	a.u.	0.9123	0.3827 - 1.0168
K_{HA_2}	P concentration of half-max P-dependent HA inhibition	a.u.	0.8323	0.372 - 1.1785
K_{HA_3}	E concentration of half-max E-dependent HA activation	a.u.	0.2352	0.0508 - 0.7787
K_{HC_1}	B concentration of half-max B-dependent HC activation	a.u.	0.2428	0.2428 - 1.1452
K_{HC_2}	N concentration of half-max N-dependent HC activation	a.u.	1.0281	0.1809 - 1.235
K_{WP_1}	N concentration of half-max N-dependent HC activation	a.u.	0.4593	0.0518 - 0.781
K_{WP_2}	E concentration of half-max E-dependent HC activation	a.u.	0.6369	0.2344 - 0.8767
K_{E_1}	HA concentration of half-max HA-dependent E activation	a.u.	0.8437	0.2917 - 1.1875
K_{E_2}	N concentration of half-max N-dependent E activation	a.u.	0.615	0.0022 - 0.615
K_{E_3}	B concentration of half-max B-dependent E activation	a.u.	0.1496	0.0025 - 0.6638
b_1	Hill coefficient HC-dependent M activation	-	0.8518	0.6624 - 0.9338
b_2	Hill coefficient P-dependent B inhibition	-	0.8224	0.7496 - 0.9762
b_3	Hill coefficient N-dependent B inhibition	-	0.7723	0.7723 - 1.2021
b_4	Hill coefficient P-dependent B inhibition	-	1.0886	0.8193 - 1.1606
b_5	Hill coefficient N-dependent B inhibition	-	0.7842	0.7214 - 0.8926
n_1	Hill coefficient P-dependent N inhibition	-	0.9602	0.7687 - 0.9602

n_2	Hill coefficient B-dependent N inhibition	-	0.842	0.7417 - 0.9579
n_3	Hill coefficient E-dependent N inhibition	-	1.1795	0.7981 - 1.2161
w_1	Hill coefficient N-dependent W activation	-	0.9507	0.7787 - 1.008
w_2	Hill coefficient P-dependent W inhibition	-	0.8503	0.7295 - 0.9602
w_3	Hill coefficient E-dependent W activation	-	1.0569	0.9993 - 1.2234
ha_1	Hill coefficient N-dependent HA activation	-	0.9471	0.6835 - 0.9847
ha_2	Hill coefficient P-dependent HA inhibition	-	0.8611	0.6277 - 0.9603
ha_3	Hill coefficient E-dependent HA activation	-	1.2213	1.1063 - 1.2342
hc_1	Hill coefficient B-dependent HC activation	-	0.9979	0.7447 - 0.9979
hc_2	Hill coefficient N-dependent HC activation	-	0.9771	0.647 - 0.9771
wp_1	Hill coefficient N-dependent HC activation	-	1.1477	1.0495 - 1.2692
wp_2	Hill coefficient E-dependent HC activation	-	1.0131	0.716 - 1.154
e_1	Hill coefficient HA-dependent E activation	-	1.1561	0.74 - 1.1561
e_2	Hill coefficient N-dependent E activation	-	1.0253	1.0253 - 1.2221
e_3	Hill coefficient B-dependent E activation	-	0.9893	0.9893 - 1.2484

Table S7. Stem cell time course model parameters

Symbol	Description	Unit	Value	Range
p_P	PAN production rate	a.u. h^{-1}	1.7339	0.2934 -2.8631
p_B	BRAVO production rate	a.u. h^{-1}	0.8572	0.5608 -3.1351
p_N	NTT production rate	a.u. h^{-1}	0.5906	0.3001 -2.7948
p_{HC}	HSFC1 production rate	a.u. h^{-1}	0.8499	0.4320 -2.7893
p_{EB}	EBP production rate	a.u. h^{-1}	2.6078	0.3783 -3.0435
p_E	EIN3 production rate	a.u. h^{-1}	3.0528	0.4209 -3.2377
p_{HA}	HCFA1E production rate	a.u. h^{-1}	2.5538	0.3671 -3.0208
p_W	WRKY23 production rate	a.u. h^{-1}	1.8797	0.4525 -3.1301

p_S	SCRM2 production rate	a.u. h^{-1}	0.2267	0.0000 -0.3157
d_P	PAN decay rate	h^{-1}	0.2192	0.0270 -0.3727
d_B	BRAVO decay rate	h^{-1}	0.0080	0.0002 -0.2513
d_N	NTT decay rate	h^{-1}	0.2178	0.0290 -0.4990
d_{HC}	HSFC1 decay rate	h^{-1}	0.1219	0.0649 -0.3863
d_{EB}	EBP decay rate	h^{-1}	0.0029	0.0009 -0.0914
d_E	EBP decay rate	h^{-1}	0.0135	0.0017 -0.1643
d_{HA}	HCFA1E decay rate	h^{-1}	0.2553	0.0110 -0.3278
d_W	WRKY23 decay rate	h^{-1}	0.4585	0.0028 -0.9770
d_S	SCRM2 decay rate	h^{-1}	0.0985	0.0024 -1.0641
K_{B1}	P concentration of half-max P-dependent B activation	a.u.	0.0679	0.0002 -1.0415
K_{B2}	EB concentration of half-max EB-dependent B inhibition	a.u.	0.1931	0.0533 -1.1721
K_{B3}	E concentration of half-max E-dependent B inhibition	a.u.	0.2894	0.0063 -1.0687
K_{N1}	P concentration of half-max P-dependent N inhibition	a.u.	0.8329	0.0695 -1.2140
K_{N2}	HC concentration of half-max HC-dependent N inhibition	a.u.	0.9804	0.1413 -1.2295
K_{N3}	W concentration of half-max W-dependent N inhibition	a.u.	0.0076	0.0001 -1.1787
K_{HC1}	P concentration of half-max P-dependent HC activation	a.u.	0.5678	0.0999 -1.2526
K_{HC2}	N concentration of half-max N-dependent HC activation	a.u.	0.9169	0.1199 -1.1949
K_{EB1}	P concentration of half-max P-dependent E activation	a.u.	0.0045	0.0001 -1.0377
K_{EB2}	B concentration of half-max B-dependent E activation	a.u.	0.1201	0.0024 -1.1947

K_{EB3}	E concentration of half-max E-dependent E activation	a.u.	0.4171	0.0097 -1.0483
K_{E1}	P concentration of half-max P-dependent E activation	a.u.	1.0881	0.0561 -1.2283
K_{E2}	N concentration of half-max N-dependent E activation	a.u.	0.4486	0.0324 -1.1502
K_{E3}	HC concentration of half-max HC-dependent E activation	a.u.	0.4713	0.0154 -1.2520
K_{HA1}	B concentration of half-max B-dependent HA activation	a.u.	0.1728	0.0419 -1.1406
K_{HA2}	E concentration of half-max E-dependent HA inhibition	a.u.	0.6195	0.0130 -0.8820
K_{W1}	S concentration of half-max S-dependent W activation	a.u.	0.1116	0.0000 -0.9446
K_{W2}	E concentration of half-max E-dependent W inhibition	a.u.	1.0999	0.0130 -1.1677
K_{S1}	B concentration of half-max B-dependent HC activation	a.u.	0.5276	0.0000 -1.0121
K_{S2}	E concentration of half-max E-dependent HC activation	a.u.	1.1531	0.8090 -1.2907
b_1	Hill coefficient P-dependent B inhibition	-	1.0223	0.7430 -1.2337
b_2	Hill coefficient EB-dependent B inhibition	-	0.8905	0.7047 -1.0973
b_3	Hill coefficient E-dependent B inhibition	-	0.8127	0.6779 -1.2322
n_1	Hill coefficient P-dependent N inhibition	-	0.8534	0.6917 -1.0528
n_2	Hill coefficient HC-dependent N inhibition	-	0.9988	0.6808 -1.2025
n_3	Hill coefficient W-dependent N inhibition	-	1.1919	0.8869 -1.3477
hc_1	Hill coefficient P-dependent HC activation	-	0.9113	0.6322 -1.1085
hc_2	Hill coefficient N-dependent HC activation	-	1.1327	0.6924 -1.1327
eb_1	Hill coefficient P-dependent E activation	-	1.1654	0.7483 -1.2756
eb_2	Hill coefficient B-dependent E activation	-	0.8445	0.7791 -1.2137
eb_3	Hill coefficient E-dependent E activation	-	0.9612	0.7994 -1.2319
e_1	Hill coefficient P-dependent E activation	-	0.7822	0.7035 -1.2289

e_2	Hill coefficient N-dependent E activation	-	1.1477	0.7366 -1.3059
e_3	Hill coefficient HC-dependent E activation	-	0.9695	0.6979 -1.2886
ha_1	Hill coefficient B-dependent HA activation	-	1.1057	0.6664 -1.1504
ha_2	Hill coefficient E-dependent HA inhibition	-	1.2065	0.7000 -1.3099
w_1	Hill coefficient S-dependent W activation	-	1.0063	0.7237 -1.2156
w_2	Hill coefficient E-dependent W inhibition	-	0.9374	0.7050 -1.2391
s_1	Hill coefficient W-dependent HC activation	-	1.0285	0.7487 -1.2681
s_2	Hill coefficient E-dependent HC activation	-	0.9486	0.7511 -1.2272

Dataset S01. ANOVA normalized gene expression values in the SCN, QC, CEI, XYL, Stages 1,2,3.

Dataset S02. Differentially expressed genes in each stem cell type

Dataset S03. Known stem cell regulators found among the differentially expressed genes

Dataset S04. Differentially expressed genes in the *pan* mutant with respect to the wild-type

Dataset S05. Inferred clusters of TFs in the network of genes enriched in the stem cells

References

1. Rousseeuw PJ. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. J Comput Appl Math. 1987;20(C):53–65.
2. Zou M, Conzen SD. A new dynamic Bayesian network (DBN) approach for identifying gene regulatory networks from time course microarray data. Bioinformatics. 2005;21(1):71–9.
3. Kim SY, Imoto S, Miyano S. Inferring gene networks from time series microarray data using dynamic Bayesian networks. Brief Bioinform. 2003;4(3):228–35.
4. Yu J, Smith VA, Wang PP. Advances to Bayesian network inference for generating causal networks from observational biological data. Bioinformatics. 2004;20(18):3594–603.
5. Buntine W. Theory refinement on Bayesian networks. In: Proceedings of the Seventh Conference on Uncertainty in Artificial Intelligence [Internet]. 1991. p. 52–60. Available from: http://dl.acm.org/citation.cfm?id=2100669
6. Shannon P, Markiel A, Ozier O, Baliga NS, Wang JT, Ramage D, et al. Cytoscape : A Software Environment for Integrated Models of Biomolecular Interaction Networks. Genome Res. 2003;13(11):2498–504.
7. Sozzani R, Cui H, Moreno-Risueno M a, Busch W, Van Norman JM, Vernoux T, et al. Spatiotemporal regulation of cell-cycle genes by SHORTROOT links patterning and growth. Nature [Internet]. 2010;466(7302):128–32. Available from: http://dx.doi.org/10.1038/nature09143
8. Nawy T, Lee J-Y, Colinas J, Wang JY, Thongrod SC, Malamy JE, et al. Transcriptional Profile of the Arabidopsis Root Quiescent Center. Dev Cell. 2005;17(July):1908–25.
9. Birnbaum K, Shasha DE, Wang JY, Jung JW, Lambert GM, Galbraith DW, et al. A gene expression map of the Arabidopsis root. Science. 2003;302(5652):1956–60.
10. Brady SM, Orlando DA, Lee J-Y, Wang JY, Koch J, Dinneny JR, et al. A high-resolution root spatiotemporal map reveals dominant expression patterns. Science [Internet]. 2007;318(5851):801–6. Available from: http://www.ncbi.nlm.nih.gov/pubmed/17975066

11. Marbach D, Prill RJ, Schaffter T, Mattiussi C, Floreano D, Stolovitzky G. Revealing strengths and weaknesses of methods for gene network inference. Proc Natl Acad Sci U S A [Internet]. 2010;107(14):6286–91. Available from: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2851985&tool=pmcentrez&rendertype=abstract
12. Marbach D, Schaffter T, Mattiussi C, Floreano D. Generating realistic in silico gene networks for performance assessment of reverse engineering methods. J Comput Biol. 2009;16:1–8.
13. Prill RJ, Marbach D, Saez-Rodriguez J, Sorger PK, Alexopoulos LG, Xue X, et al. Towards a rigorous assessment of systems biology models: The DREAM3 challenges. PLoS One. 2010;5(2).
14. Rau A, Jaffrézic F, Foulley J-L, Doerge RW. An empirical Bayesian method for estimating biological networks from temporal microarray data. Stat Appl Genet Mol Biol [Internet]. 2010;9(1):Article 9. Available from: http://dx.doi.org/10.2202/1544-6115.1513
15. Young WC, Raftery AE, Yeung KY. Fast Bayesian inference for gene regulatory networks using ScanBMA. BMC Syst Biol [Internet]. 2014;8(1):47. Available from: http://www.ncbi.nlm.nih.gov/pubmed/24742092
16. Margolin AA, Nemenman I, Basso K, Wiggins C, Stolovitzky G, Dalla Favera R, et al. ARACNE: an algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context. BMC Bioinformatics. 2006;7 Suppl 1:S7.
17. Faith JJ, Hayete B, Thaden JT, Mogno I, Wierzbowski J, Cottarel G, et al. Large-scale mapping and validation of Escherichia coli transcriptional regulation from a compendium of expression profiles. PLoS Biol. 2007;5(1):0054–66.
18. Meyer PE, Kontos K, Lafitte F, Bontempi G. Information-theoretic inference of large transcriptional regulatory networks. Eurasip J Bioinforma Syst Biol. 2007;2007.
19. Tibshirani R. Regression Selection and Shrinkage via the Lasso. J R Stat Soc B [Internet]. 1994;58(1):267–88. Available from: http://citeseer.ist.psu.edu/viewdoc/summary?doi=10.1.1.35.7574
20. Tibshirani R, Saunders M, Rosset S, Zhu J, Knight K. Sparsity and smoothness via the fused lasso. J R Stat Soc Ser B Stat Methodol. 2005;67(1):91–108.
21. Furuta KM, Miyashima S, Rybel B De, Somervuo P, Rocha R, Thitamadee S, et al. Arabidopsis NAC45 / 86 direct sieve element morphogenesis culminating in enucleation. 2014;(July):1–7.
22. Vermeirssen V, De Clercq I, Van Parys T, Van Breusegem F, Van de Peer Y. Arabidopsis\n Ensemble Reverse-Engineered Gene Regulatory Network Discloses Interconnected Transcription Factors in Oxidative Stress. Plant Cell Online [Internet]. 2014;26(12):4656–79. Available from: http://www.plantcell.org/lookup/doi/10.1105/tpc.114.131417
23. Hallinan JS, Jackway PT. Network Motifs, Feedback Loops and the Dynamics of Genetic Regulatory Networks. 2005 IEEE Symp Comput Intell Bioinforma Comput Biol. 2005;
24. Orrell D, Bolouri H. Control of internal and external noise in genetic regulatory networks. J Theor Biol. 2004;230(3):301–12.
25. Taylor-Teeples M, Lin L, de Lucas M, Turco G, Toal TW, Gaudinier A, et al. An Arabidopsis gene regulatory network for secondary cell wall synthesis. Nature. Nature Publishing Group; 2015;517(7536):571–5.
26. Shen-Orr SS, Milo R, Mangan S, Alon U. Network motifs in the transcriptional regulation network of Escherichia coli. Nat Genet [Internet]. 2002;31(1):64–8. Available from: http://www.nature.com/ng/journal/v31/n1/full/ng881.html
27. Albert R, Barabasi AL. Statistical mechanics of complex networks. Rev Mod Phys [Internet]. 2002;74(1):47–97. Available from: http://dx.doi.org/10.1103/RevModPhys.74.47%5Cnpapers2://publication/doi/10.1103/RevModPhys.74.47
28. Albert R. Scale-free networks in cell biology. J Cell Sci. 2005;118(Pt 21):4947–57.
29. Benfey PN, Linstead PJ, Roberts K, Schiefelbein JW, Hauser M-T, Aeschbacher RA. Root development in *Arabidopsis*: four mutants with dramatically altered root morphogenesis. Development [Internet]. 1993;119:57–70. Available from: http://dev.biologists.org/cgi/content/abstract/119/1/57
30. Cruz-Ramirez A, Diaz-Triviño S, Blilou I, Grieneisen VA, Sozzani R, Zamioudis C, et al. A bistable circuit involving SCARECROW-RETINOBLASTOMA integrates cues to inform asymmetric stem cell division. Cell. 2012;150(5):1002–15.

31. Sabatini S, Heidstra R, Wildwater M, Scheres B. SCARECROW is involved in positioning the stem cell niche in the Arabidopsis root meristem. Genes Dev. 2003;17(3):354–8.
32. Levesque MP, Vernoux T, Busch W, Cui H, Wang JY, Blilou I, et al. Whole-genome analysis of the short-root developmental pathway in Arabidopsis. PLoS Biol. 2006;4(5):739–52.
33. Welch D, Hassan H, Blilou I, Immink R, Heidstra R, Scheres B. Arabidopsis JACKDAW and MAGPIE zinc finger proteins delimit asymmetric cell division and stabilize tissue boundaries by restricting SHORT-ROOT action. Genes Dev. 2007;21(17):2196–204.
34. Helariutta Y, Fukaki H, Wysocka-Diller J, Nakajima K, Jung J, Sena G, et al. The SHORT-ROOT gene controls radial patterning of the Arabidopsis root through radial signaling. Cell [Internet]. 2000;101(5):555–67. Available from: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=10850497&retmode=ref&cmd=prlinks%5Cnpapers3://publication/doi/10.1016/S0092-8674(00)80865-X
35. Galinha C, Hofhuis H, Luijten M, Willemsen V, Blilou I, Heidstra R, et al. PLETHORA proteins as dose-dependent master regulators of Arabidopsis root development. Nature. 2007;449(7165):1053–7.